RNA BIOLOGY
2020, VOL. 17, NO. 11, 1657–1665
https://doi.org/10.1080/15476286.2020.1792116

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER



# Comparative analysis of long noncoding RNAs in long-lived mammals provides insights into natural cancer-resistance

Jian-Jun Jiang[a,b] and Qing-Peng Kong[a,c,d,e]

[a]State Key Laboratory of Genetic Resources and Evolution/Key Laboratory of Healthy Aging Research of Yunnan Province, Kunming Institute of Zoology, The Chinese Academy of Sciences, Kunming, China; [b]Kunming College of Life Science, University of Chinese Academy of Sciences, Beijing, China; [c]CAS Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China; [d]Kunming Key Laboratory of Healthy Aging Study, Kunming, China; [e]KIZ/CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming, China

**ABSTRACT**

Mouse and rats are staple model organisms that have been traditionally used for oncological studies; however, their short lifespan and highly prone to cancers limit their utilizationsin understanding the mechanisms of cancer resistance. In recent years, several studies of the non-standard long-lived mammalian species like naked mole rat (NMR) have provided new insights of mechanisms in natural anti-cancer. How long-lived species genetically maintain longevity and cancer-resistance remains largely elusive. To better understand the underlying anti-cancer mechanisms in long-lived mammals, we genome widely identified long noncoding RNA (lncRNA) transcripts of two longevous mammals, bowhead whale (BW, *Balaena mysticetus*) and Brandt's bat (BB, *Myotis brandtii*) and featured their sequence traits, expression patterns, and their correlations with cancer-resistance. Similar with naked mole rat (NMR, *Heterocephalus glaber*), the most long-lived rodent, BW and BB lncRNAs show low sequence conservation and dynamic expressions among tissues and physiological stages. By utilizing k-mers clustering, 75–136 of BW, BB and NMR lncRNAs were found in close relation (Pearson's $r \geq 0.9$, $p < 0.01$) with human ageing diseases related lncRNAs (HAR-Lncs). In addition, we observed thousands of BB and BW lncRNAs strongly co-expressed ($r > 0.8$ or $r < -0.8$, $p < 0.01$) with potential tumour suppressors, indicating that lncRNAs are potentially involved in anti-cancer regulation in long-lived mammals. Our study provides the basis for lncRNA researches in perspectives of evolution and anti-cancer studies.

**Abbreviations:** BW: bowhead whale; BB: Brandt's bat; NMR: naked mole rat; LLM: long-lived mammal; HTS: human tumour-suppressors; PTS: potential tumour suppressor; ARD: ageing related diseases; HAR-Lncs: lncRNAs that related with human ageing diseases; Kmer-lncs: lncRNAs in long-lived mammal species that corelated (Pearson's$r \geq 0.9$, $p < 0.01$) with the 10 HAR-Lncs by k-mers clustering; All-lncs: all the lncRNAs in long-lived mammal species; SDE-lncs: significant differentially expressed lncRNAs

**ARTICLE HISTORY**
Received 17 February 2020
Revised 6 June 2020
Accepted 30 June 2020

**KEYWORDS**
Long noncoding RNA; bowhead whale; Brandt's bat; naked mole rat; longevity; cancer-resistance

## Introduction

Cancer is the most common cause of human death worldwide.In the past decades, mouse and rat as the traditional model organisms have been widely used in cancer studies. The long-term goal of cancer research is to develop a therapeutic or preventive strategy with both efficiency and non-toxicity [1,2]. However, more than 50% of aged mice and rats in previous studies die of cancer while this number in humans is approximately 23% [3], suggesting staple mammal models are limited for understanding and exploring the spontaneous occurring anti-cancer strategies. The cancer rates and ages of disease onset inmammal species have diverged dramatically during the lengthy process of evolution [1].There are several species include the naked mole rat, blind mole rat, Brandt's bat, elephant and bowhead whales are known for their extremely longevity and cancer resistance [4]. It takes averagely two years for a mouse to develop cancer, whereas it takes decades in long- lived species [1,5–7], suggesting that longevous species are good animal models for studying the natural anti-cancer mechanisms.

Researchers have found multiple mechanisms contribute to the remarkable cancer resistance of naked mole rat (*Heterocephalus glaber*), the longest-live rodent. The cells of the naked mole rat secreted unique high molecular mass hyaluronan (HMM-HA) resulting in the activation of early contact inhibition (ECI) while cancerous cells lose ECI and continue to proliferate on top of each other [8]. Inactivation of either one of the tumour suppressors like *RB1* or *TP53* causes cell apoptosis in naked mole rat; however, it is in strikingly contrast to mouse or human cells, where inactivationof either *RB1* or *TP53* leads to more rapid proliferation [9].Bats live longer than any other order of mammal relative to their body mass [10]. Therefore, it is surprising that only a few incidences of tumours have been reported in bats [1]. In particular, the Brandt's bat (*Myotis brandtii*) holds the longevity record with a small size of approximate 7 grams among

**CONTACT** Qing-Peng Kong kongqp@mail.kiz.ac.cn State Key Laboratory of Genetic Resources and Evolution/Key Laboratory of Healthy Aging Research of Yunnan Province, Kunming Institute of Zoology, The Chinese Academy of Sciences, Kunming 650223, China
Supplemental data for this article can be accessed here.

the bats [11]. Recent studies suggested that an up-regulation of tumour suppressor miRNAs and the distinctive mitochondrial function in bats have played critical roles in their cancer-resistance [12–15]. On the other hand, the genetic variations ofthe highly conserved transmembrane domains in the growth hormone receptor (GHR) and insulin-like growth factor 1 receptor (IGF1 R)of long-lived bats are associated factors that contribute to the cancer resistance [13]. In addition to small sized long-lived mammals, bowhead whale (*Balaena Mysticetus*), the most longevous mammal species with large body size has evolved extraordinary tumor suppressor mechanisms as well. Previously comparative genomic and transcriptomic studies found the specific changes of gene expression include insulin pathway genes and some positive selected genes like *ERCC1, PCNA* and *UCP1* in bowhead whale are linked to cancer and ageing [16,17].

Accumulated evidence has underlined the protein coding genes (PCG)-based genetic mechanisms of cancer-resistance in long-live mammals; however, the roles of epigenetic factors like long noncoding RNA (lncRNA) are remain largely unknown. The advance and accuracy of sequencing technology in recent years made lncRNAs the potential diagnose and prognostic markers in human cancers. Large number of lncRNAs were dysregulated in various cancer types, and many of them have been identified as tumour drivers such as *LINK-A* [18,19] or potential suppressors like *Gas5* [20]. In human triple-negative breast cancer (TNBC), an upregulation of *LINK-A* was particularly observed comparing with non-TNBC breast cancer tissues and it is predictive of poor prognosis for patients with breast cancer [18]. In contrast to *LINK-A*, the expression of tumour suppressor, *GAS5*, is typically reduced in multiple cancers [21]. More crucially, the low *GAS5* lncRNA levels are correlated with poor outcomes [22,23]. Recent studies revealed that overexpression of *GAS5* can significantly suppress the growth of tumour cells in many cancer types [24–26], demonstrating that lncRNAs are promising therapeutic strategies for cancer treatment. For a better knowledge of lncRNAs in anti-cancer mechanisms, we identified the high-quality lncRNAs in two non-standard long-lived mammals organisms, Brandt's bat and bowhead whales, at the genome wide scale; and performed comparative analysis with the published naked mole rat lncRNAs, aiming at mining the expressing features of long-lived mammals lncRNAs as well as their potential associations with natural cancer-resistance.

## Results

### *Genome-wide identification of lncRNAs in long-lived mammals (LLM)*

We genome-widely identified lncRNAs intwo long-lived species include bowhead whale (BW) and Brandt's bat (BB). RNA-seq profiles containing 6 BW transcriptomes and 15 BB transcriptomes were used for lncRNAs identification (Table S1).

Following a stringent lncRNA identification pipeline, we separately annotated 8,004 and 6,824 high-quality lncRNAs transcripts from BW and BB genomes. To better understand the associations among lncRNAs, longevity and cancer-resistance, 4,422 lncRNA sequences from the most longevous rodent – naked mole rat (NMR) were acquired from our previous study for comparative analysis [27]. In comparison with protein coding gene (PCG), lncRNA in all three longevous species exhibit significant differences, with lower average lengths but higher GCcontents (Fig. 1). Average length of LLM lncRNAs is significant shorter than PCG ($p < 0.01$), especially in BB and NMR lncRNAs. Comparing with the

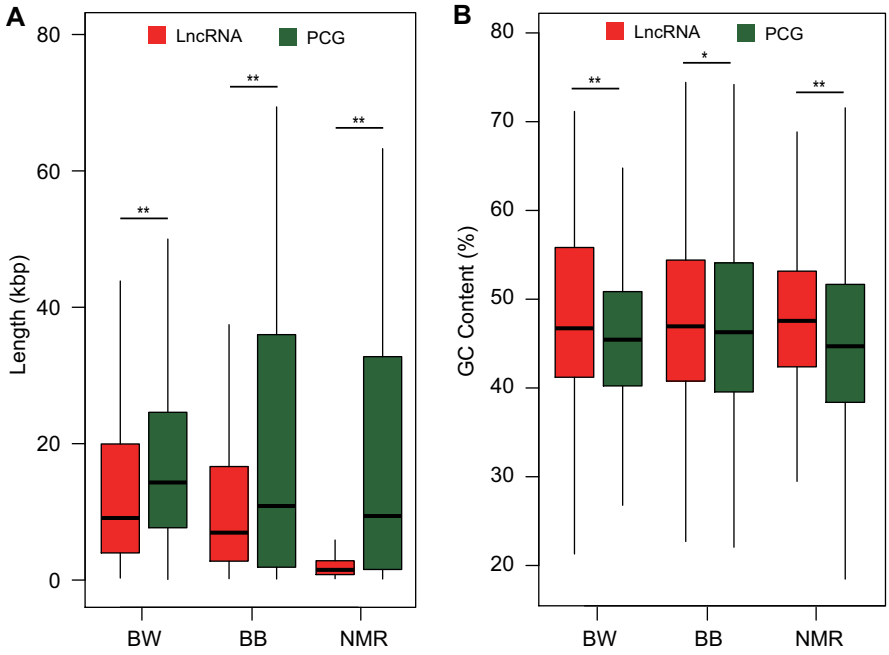


**Figure 1.** Features of LLM lncRNAs comparing with protein coding genes (PCGs). (A) Average Length. (B) Average GC contents.

other two longevous species, NMR possesses the shortest lncRNAs at average.

### Conservation analysis among LLM lncRNAs

As suggested by previous phylogenic studies, the closest relative to the BW is minke whale (*Balaenoptera acutorstrata*), their divergence time is estimated by 25–30 million years (MYA) ago [17,28]. The life expectation of BB is 200 years; however, it is only 50 years for minke whale (Fig. 2). NMR has alike look and body size with mouse except for its barely nude skin, it is estimated that NMR divergent from rodent around 73 (60.2–80.3) MYA [29], and their life spans differ to 10 folds. To explore the conserved lncRNAs among these three longevous species, orthologous analysis was performed using OrthoMCL with BLASTN programme[30]. Only 2 pairs of BB and BW lncRNAs show homology (≥ 50% hits and E-value<$1e^{-5}$), suggesting that LLM lncRNAs lack sequence conservation. The homologous lncRNA information are listed: (1) BW_00107607|location: scaffold_3048:284–646 *vs*BB_00066642|location: scaffold4:7,536,925–7,537,412; (2) BW_00071170|location: scaffold_21:1,920,428–1,923,505 *vs* BB_00072718|location: scaffold454:1,106,022–110,927.

By analysing their FPKM values, we found that the first pair of homologs are completely non-detectable in all tissues except brain (cerebrum) in both BB and BW; however, the other pair of lncRNAs show less tissue-specific but express significant higher in brains as well. The functions of those homologous lncRNAs are currently unknown, but their high or unique expression in brains suggested that they may involve in neuronal related functionality.

### Spatiality- and tissues-specific expression of lncRNAs in BB and BW genomes

The 6 BW transcriptomes were isolated and prepared from 6 tissues consisting of liver, cerebellum, heart, lung, kidney and muscle from a male bowhead whale [17]. A total of 15 RNA-seq profiles of BBwere obtained from 3 tissues (liver, kidney and brain) in bats at three behavioural periods include hibernating for 2 months (H2), hibernating for 6 months (H6) and summer active (A) [11].

Behavioural studies indicate someecological and physiological features in bats like hibernationcan increase their survival rates, thus linking with the longevity [31,32]. During hibernating, hundreds of lncRNAs in BB expressed differentially among tissues especially in livers (Fig. 3B–D). Comparing with summer active bats, there are 843 and 441 lncRNAs dysregulated in liver tissues of H2 and H6 bats, respectively; of which 229 are overlapped (Fig. 3C). In addition, 62 BB lncRNAs expressed significantly different acrossthree periods in livers (Fig. 3C); and a total of 26 dysregulated lncRNAs were commonly found in three kidneys (Fig. 3D).Similar with BB lncRNAs, 661 BW lncRNAs distribute specific expression manners across 6 tissues and the top-ranked 50 lncRNAs are shown in Fig. 3A.The expression of BB and BW lncRNAs displayed spatiality- and tissues- specificity, which is in common with NMR and non-longevous species [27].

### Co-expression analysis among BB lncRNAs, BW lncRNAs and tumour suppressors

To investigate the associations between LLM lncRNAs and cancer resistance, human tumour suppressing genes were utilized for co-expression analysis. We organized a gene list that contains 1,217 experimentally verified human tumour-suppressors (HTSs) from TSGene2 [33], and then searched their orthologs in BB and BW from the annotated genomic files. In BW genome, 940 PCGs were identified as homologs of HTSs, and it was 1,015 in BB species. These homologs were considered as potential tumour suppressors in BB (PTSs-BB) and BW (PTSs-BW). Co-expression analysis of LLM lncRNAs and PTSs-LLM was performed by in-house R script.

The expression of lncRNA is usually at low level, which will probably result in some false positive results in co-expression analysis. To eliminate the bias, we only retained lncRNAs that expressed in all the tissues with FPKM values ≥0.3. In this way, a total of 3,961 BW lncRNAs and 4,070 BB lncRNAs were used for the subsequent analyses. Surprisingly and interestingly, both Pearson and Spearman correlation ranking analysis indicate that 89.7% (3,554/3961) BW lncRNAs show strong co-expression($r > 0.8$ or $r < -0.8$, $p < 0.01$) with 646 PTSs-BW. For BB species, Spearman ranking exhibits a similar but tiny higher rate (99%, 4032/4070) of BB lncRNAs that strongly (*r*

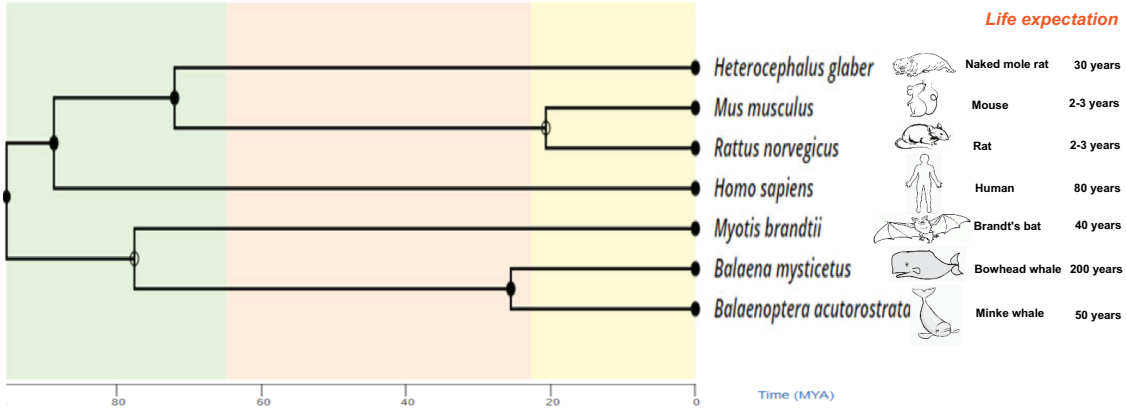


**Figure 2.** Phylogenetic tree of long-lived mammals and their relative species.

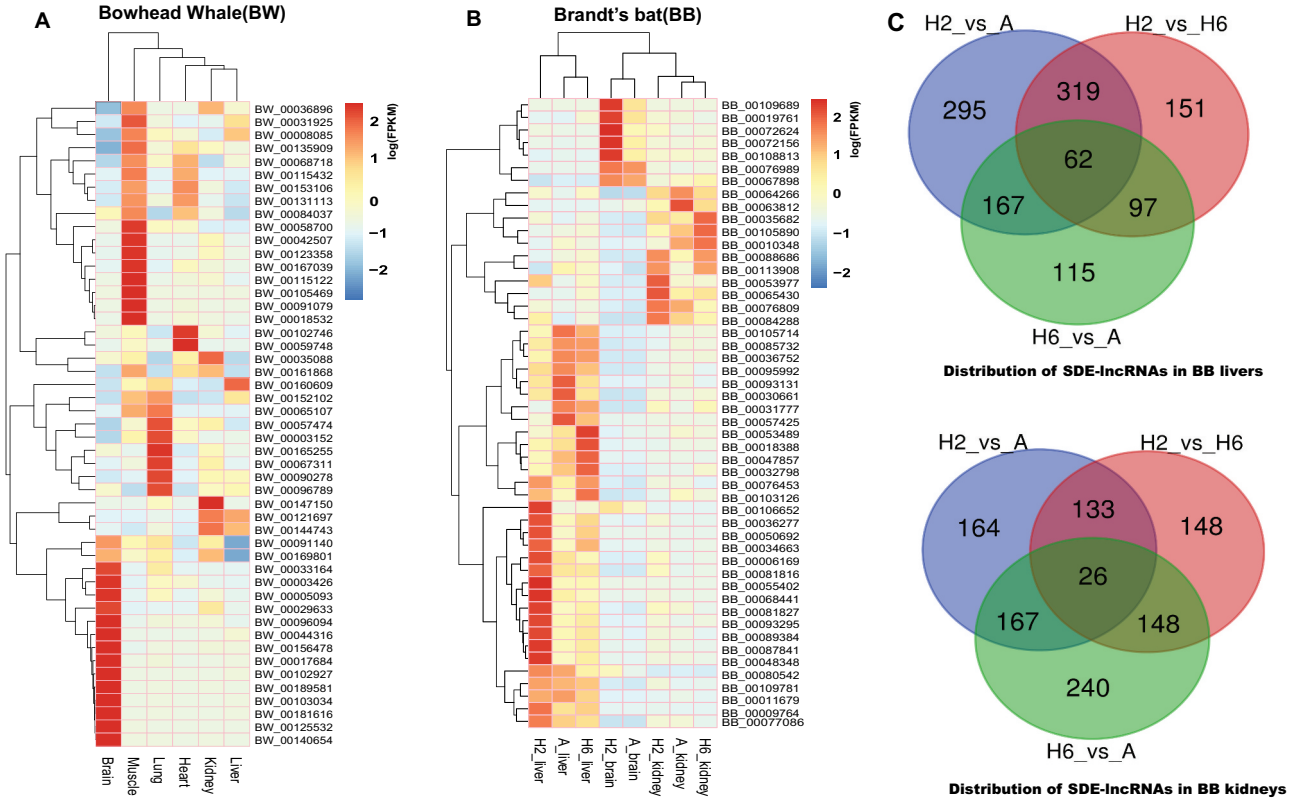


**Figure 3.** Significant differentially expressed (SDE) -lncRNAs in BB and BW genomes among tissues and physiological stages ('A' means summer active bats; 'H2' and 'H6' separately meansbats hibernate for 2 and 6 months). (A) Heatmap of the top 50 BWSDE-lncRNAs across 6 tissues (B) Heatmap of the top 50 BB SDE-lncRNAsacross 3 tissues at 3 physiological stages ('A', 'H2' and 'H6') (C) Distribution of BB SDE-lncRNAsin livers (upper panel) and kidneys (lower panel) across 3 physiological stages(H2_*vs*_A: comparing 'H2' with 'A'; H2_*vs*_H6: comparing 'H2' with 'H6'; H6_*vs*_A: comparing 'H6' with 'A') by Venny diagram: in livers, 62 SDE-lncRNAs are overlapped among three pairs, 295 specific SDE-lncRNAs in H2_*vs*_A, 151 specific SDE-lncRNAs in H2_*vs*_H6, and 115 SDE-lncRNAs in H6_ *vs*_A; in kidneys, 26 SDE-lncRNAs are overlapped among three pairs, 164 specific SDE-lncRNAs in H2_*vs*_A, 148 specific SDE-lncRNAs in H2_*vs*_H6, and 240 SDE-lncRNAs in H6_ *vs*_A.

> 0.8 or $r < -0.8$, $p < 0.01$) co-expressed with 640 PTSs-BB than Pearson's (98.9%, 4027/4070) (Table S2).

Recent studies found the additional copies of TP53 in elephants genome contribute to their low risk of cancers [34]. However, in the genomes of BB, BW and NMR, non-similar extra copies of TP53 was observed. Here we found 106 BB and 432 BW lncRNAs strikingly coexpressed with *TP53* and/or *TP53INP1* genes(Pearson's $r \geq 0.9$, $p < 0.01$). InNMR, 133 lncRNA display striking co-expression with *TP53, TP53BP1* and/or *TP53INP1* genes (Pearson's $r \geq 0.9$, $p < 0.01$) (Table S2) [27]. Our co-expression analysis indicates that there exists large number of lncRNAs in long-lived mammals may relate with their natural anti-cancer mechanisms.

## *Kmers-analysis of LLM lncRNAs with human ageing diseases related lncRNAs (HAR-Lncs)*

Unlike PCGs, lncRNAs usually lack linear sequence homology despite of similar functions, which makes it difficult to predict their potential function. Recently, Kirk et al developed a sequence comparison method to deconstruct linear sequence relationships in lncRNAs and evaluate similarity based on the abundance of short motifs called k-mers, and they found that lncRNAs of related functions often had the similar k-mers contents [35].

LLMs have been reported to have low incidences of ageing related diseases(ARD) [17,36,37]. To explore possible links that connect LLM lncRNAs and their anti-ARD, we utilized k-mers clustering method to identify LLM lncRNAs that are potentially related with ARD, we collected a list with 29 humanlncRNAs (HAR-Lncs) [38] like *Hotair* and *Malat1* that have been experimental validated to be ARD related especially with cancers (Table S3). Considering those lncRNAs act as multiple roles in ARD and ageing process, we firstly performed a classification on 29 HAR-Lncs by k-mers clustering (Fig. 4). HAR-Lncs were clearly clustered into 7 modules, indicating that HAR-Lncs have 7 types of k-mers distributions (Fig. 4). The biggest HAR-Lnc module consists of 10 lncRNAs (Table 1). We subsequently synthesized the k-mers profiles of all the LLM lncRNAs and compared them with the k-mers contents of the biggest HAR-Lnc module. By calculating the Pearson correlation coefficient (Pearson's *r*), 96 of NMR lncRNAs, 136 of BB lncRNAs and 75 of BW lncRNAs displayed extremely high correlation (Pearson's $r \geq 0.9$, $p < 0.01$) with the 10 HAR-Lncs (Fig. 5, Table S4). In order to illuminate lncRNAs of other interested species that potentially regulating ageing diseases, we generated the k-mers libraries of human and mouse lncRNAs. Using the same parameters, 169 of human lncRNAs and 153 of mouse lncRNAs are highly clustered (Pearson's $r \geq 0.9$, $p < 0.01$) with HAR-Lncs, which is much more than we identified in longevous species (Table S4).

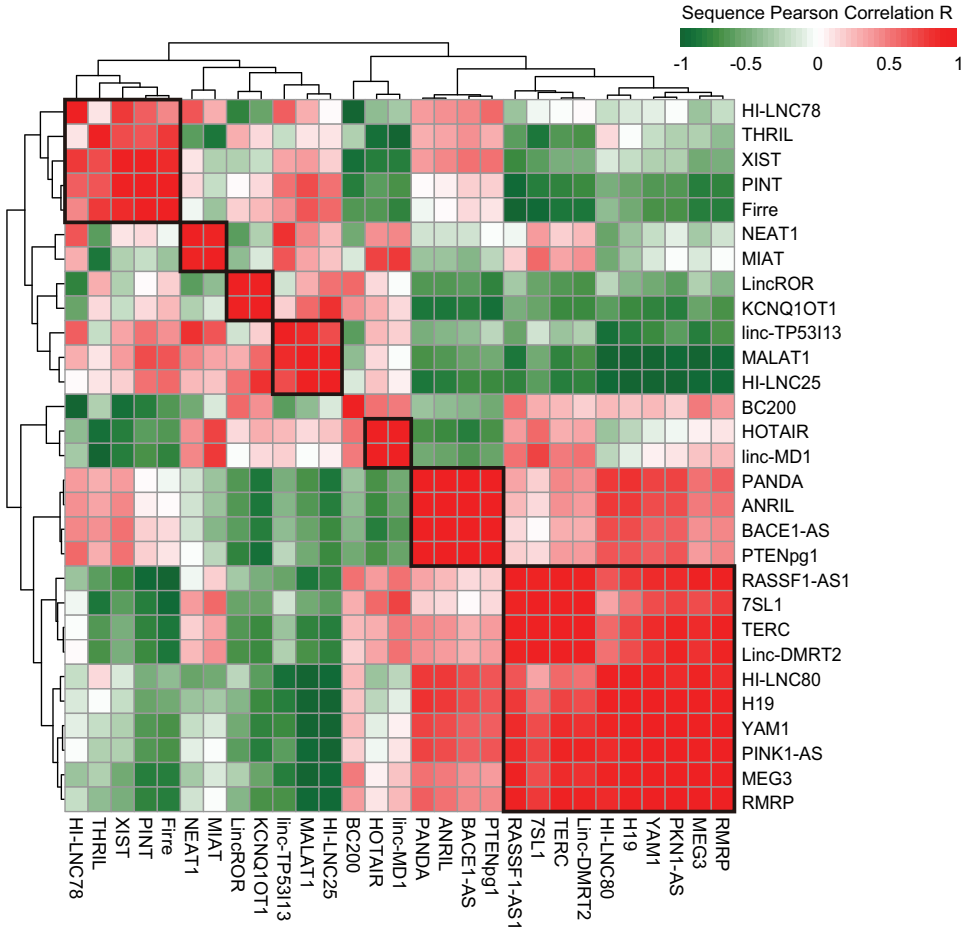


**Figure 4.** 29 HAR-Lncs are clustered into 7 blocks by k-mers distribution: the biggest block contains 9 lncRNAs.

**Table 1.** Description ofthe 10 HAR-Lncs that clustered into the biggest moduleby k-mers distribution.

| LncRNA | Relevance to human ageing diseases/process |
|---|---|
| *TERC* | RNA component of telomerase involved in telomere maintenance |
| *Linc-DMRT2* | Low in fat tissue of obese humans |
| *HI-LNC80* | Elevated by high glucose |
| *H19* | High expressed in old muscle |
| *YAM1* | Implicated in muscle wasting, cancer, diabetes, chronic heart failure |
| *PINK1-AS* | Inactive muscle |
| *RASSF1-AS1* | Involved in p53 mediated cell cycle regulation and apoptosis |
| *7SL1* | Regulates p53 expression and ER mediated transport |
| *MEG3* | PAH-related vascular remodelling by p53 pathway |
| *RMRP* | Mitochondria functional |

The 10 HAR-Lncs in the biggest module have been found to involve in tumorigenesis via p53 pathway or regulating metabolic process. For example, *TERC* is the telomerase RNA component which plays the decisive role of replicate senescence through telomere maintenance [39]. The telomerase activity in normal human somatic cells is either non-expressed or exhibiting at low levels whereas it is widely detectable in more than 85% of human tumours [40] . H19 is an imprinted oncofetal lncRNA that expressed in the embryo, down regulated at birth and then reappears in tumours and accumulated evidence has suggested its key function as an oncogene in multiple cancers [41]. In order to investigate the potential effects of the LLM lncRNAs that clustered with 10 HAR-Lncs (Kmer-lncs), we compared their average expression levels with it of all the LLM lncRNAs (All-lnc) by assessing FPKM values. In BB, Kmer-lncs expressed higher in brains and kidneys in both summer active (A) and hibernating periods (H2 and H6) but exhibited no difference in liver tissues (Fig. 6A). Similar with BB, the Kmer-lncs of BW averagely expressed higher in brain and kidneys than All-lncs (Fig. 6B). On the contrary with BB and BW, a lower expression level was observed in NMR Kmer-lncs especially in kidney tissues at three development stages that include 4-years-old (4y), 20-years-old (20y) and 4-years-old with low-oxygen-treated (LO) of naked mole rat (Fig. 6C). The function of the Kmer-lncs remains unknown, but our clustering analysis suggested that a lot of lncRNAs in the long-lived mammals are closely correlated with human cancer related lncRNAs in k-mers distribution and their dysregulated expression in tissues may potentially involve in natural cancer-resistance in long-lived mammals.

## Discussion

As the most longevous mammal, bowhead whale is always a mysterious cetacean with weight up to 100 tons. It has been proposed that the longevity of giant mammals, bowhead whale as well as elephant, stem from a lack of non-human predators and slow development [16]. Nevertheless, a specific

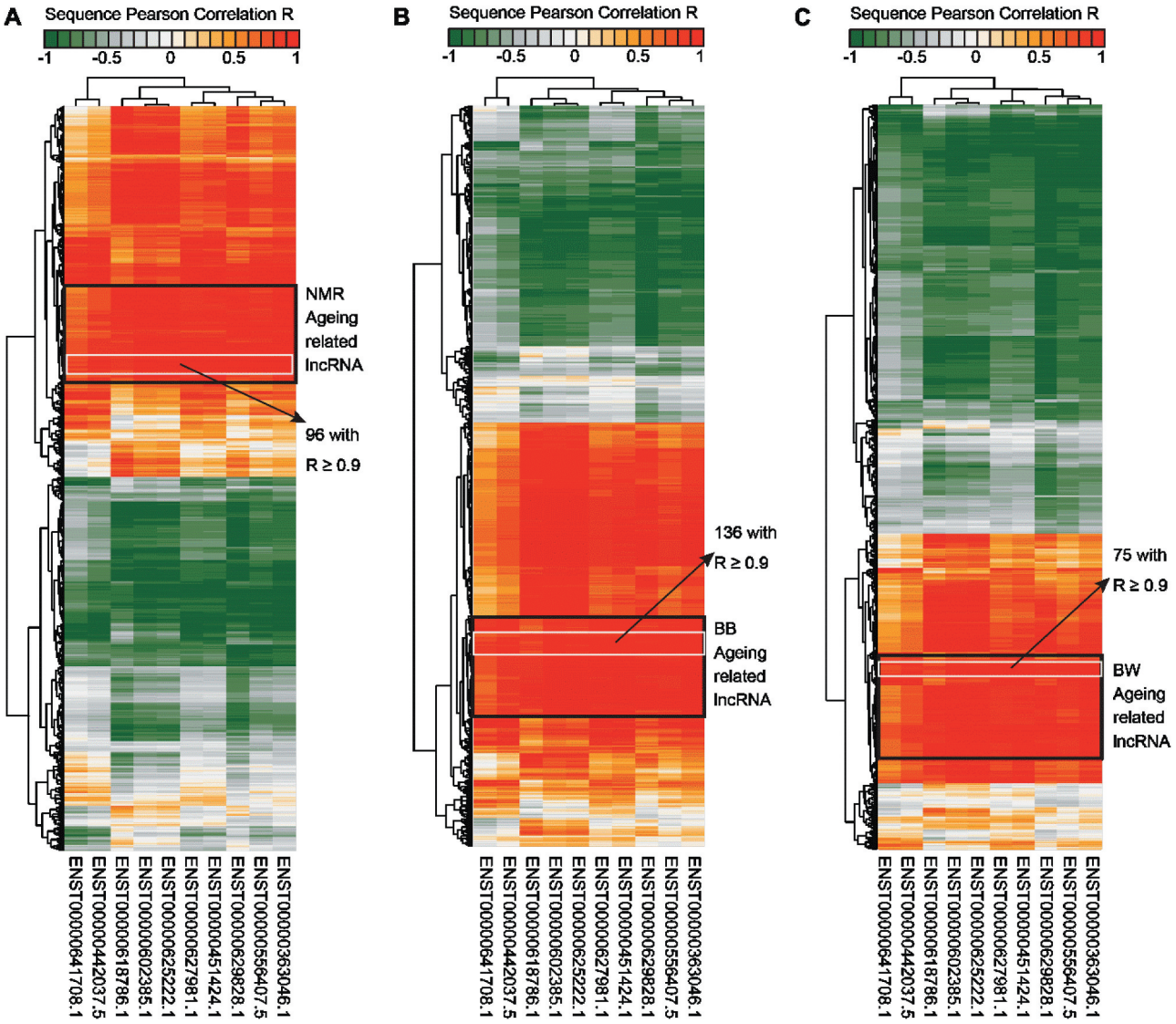


**Figure 5.** Heatmap of the correlation between LLM lncRNAs and thebiggest k-mer block of HAR-Lncs: a total of 75–136 of LLM lncRNAs show strikingly correlation (Pearson's $r \geq 0.9$, $p < 0.01$) with these 9 HAR-Lncs. (A) NMR. (B) BB. (C) BW.

expansion of *TP53* copies in elephant genome has been associated with the evolution of increased body size and an enhanced DNA damage [34,42], which potentially delays their senescence and avoids cancers. Non-similar extra copies of *TP53* was observed in bowhead whale, therefore, its underlying mechanisms that behind the cancer-resistance and longevity are still in urgent need to be recognized. Unlike bowhead whale and elephant, naked mole rat and Brandt's bat are excluded from the canonical ageing theory of 'bigger animals live longer' [43] with surviving exceeds to > 30 years and a mouse-like body size. Natural selection favours them slow ageing, ageing-tolerance and cancer-resistance; however, the genetic mechanisms are rarely explored.

In recent years, lncRNAs have emerged as key modulators in regulating cancer and anti-cancer mechanisms and have been implicated as putative effective therapeutic strategies for cancer treatment [44]. However, the functions of the vast majority of lncRNAs are unknown due to very limited methods include co-expression analysis and k-mers clustering can used for their functional predictions. In order to understand the lncRNAs in non-standard longevous mammal organisms, here we assembled and identified thousands of high-quality lncRNA transcripts in two long-lived mammals, bowhead whale and Brandt's bat, and performed the comparative analysis with naked mole rat to seek for the potential associations between lncRNAs and their natural cancer-resistance. The lncRNA transcripts that we identified in this study provide basic data for future ncRNA evolutionary study. We also observed that LLM lncRNAs similar with other non-longevous species, exhibit dynamic expression patterns spanningacross tissues and different behavioural stages. Using k-mers clustering and co-expression methods, wefoundthousands of them are strikingly co-expressed with tumour suppressors and identified hundreds themshare similarities with human ageing diseases related lncRNAs (HAR-Lncs).

Although lncRNA usually lacks canonical sequence conservation like protein coding gene does across species, accumulative evidence indicated that many lncRNAs are functional conserved [45,46]. Some functional conserved

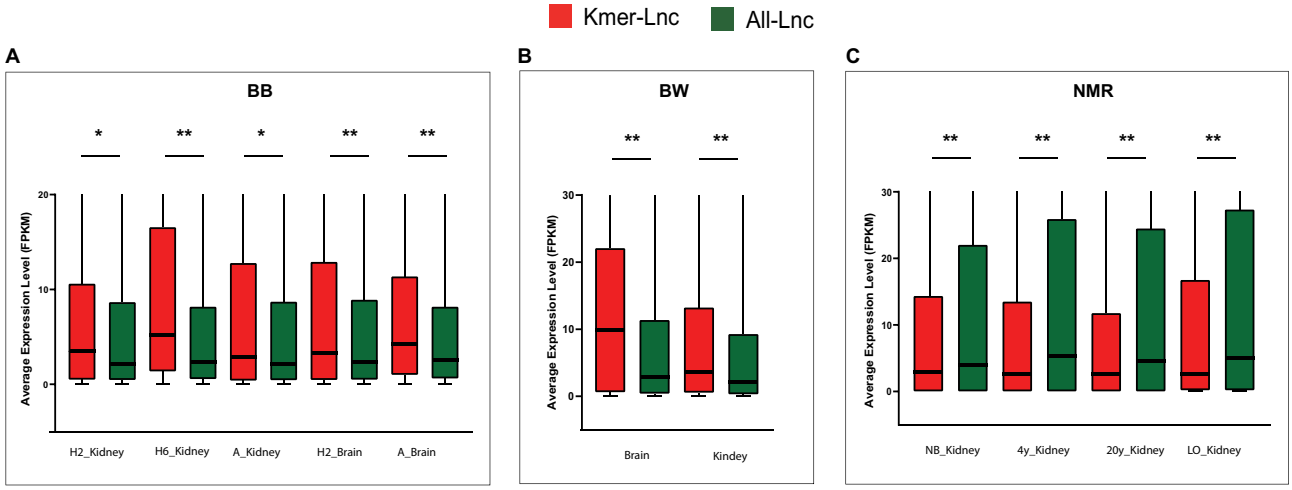


**Figure 6.** Average expression of Kmer-lncs and All-lncs show significant difference (* indicates $p < 0.05$, ** indicates $p < 0.01$) in brains and kidneys at specific stages/ages of three LLMs. (A) BB ('A': summer active bats; 'H2' and 'H6' separately means bats hibernate for 2 and 6 months, Kmer-lncs). (B) BW. (C) NMR ('NB': new-born NMR; '4y' and '20y': NMR at 4-years old and 20-years old;'LO': 4-years old NMR with low oxygen treated).

lncRNAs like *LINC-PINT* genes in human and mouse, harbour short sequence similarities that confer shared functions, even if conventional alignment algorithms do not detect the similarity [35,47]. We currently do not know whether there exists any anti-cancer related lncRNAs in long-lived mammals are playing conserved roles in human or not, the aim that we performed lncRNA identification and associated analysis is to provide basis for future lncRNA-related cancer preventive studies and trials. In summary, our cross-species investigation provides the first comprehensive genome-wide analysis of LLMs lncRNAs and reveals their prospective associations with cancer resistance, which is of significance for a better understanding of human anti-cancer study.

## Materials & methods

### Data accessibility

The raw BW and BB transcriptome datasets were downloaded from *National Centre for Biotechnology Information* (NCBI) database (PRJNA178678 and PRJNA194091). The assembled genome sequence and genomic annotation files were requested from GIGA database (http://gigadb.org/dataset/100022) and bowhead whale genome resources database (http://www.bowhead-whale.org/). NMR profiles were obtained from our previous study [27] and the naked mole rat genome resources database (http://www.naked-mole-rat.org/). Human (v29) and mouse (M19) lncRNA sequences were acquired from GENCODE database.

The fasta files, FPKM values and co-expression files of BW, BB and NMR lncRNAs by Pearson and Spearman methods can be accessed via:https://onedrive.live.com/?id=189757ACF263B018%21106&cid=189757ACF263B018, or https://drive.google.com/drive/folders/1-A2aXZgU8a8OGeSilf7_3YkC0EdFNt05?usp=sharingor clicking this link: https://pan.baidu.com/s/1e46odTZnr7H83IuBZalYJw with code: 'yvx3'

### LncRNA identification pipeline

A stepwise filtering workflow that based on our previous pipeline with minor revisions was used to identify long noncoding RNAs in two longevous mammals [27]. Briefly, raw RNA-seq reads were firstly trimmed by SolexaQA [48]. The cleaned data were then mapped to genome using Hisat2 [49] and assembled by Cufflinks packages [50]. Assembled and merged transcripts were filtered by length. Transcripts that longer than 200bp were retained for assessing their coding potential by Coding Potential Calculator 2 (CPC2) [51]. Transcripts that labelled with 'noncoding RNA' by CPC2 and their exon numbers larger than 2 were defined as long noncoding RNA.

### Conservation analyses

OrthoMCL package-BLASTN programme [30] was employed to analysis the sequence conservation among LLM lncRNAs. BLASTN was selected and transcripts with ≥ 50% coverage and E-value < 1e-5 were clustered as homologous groups.

### Functional prediction by K-mer analysis

SEEKR packages were obtained from GitHub [35] to synthesize the k-mers profiles and to calculate the similarities among lncRNAs from different species following the tutorials with default parameters.

### Co-expression analysis

To investigate the potential correlations between lncRNAs and cancer-resistance in LLM, we performed co-expression analysis between LLM lncRNAs and PTSs-LLM by testing their FPKM values. LLM lncRNAs and PTSs that expressed in all tissues with FPKM ≥ 0.3 were used for the analysis. Cufflinks package was used to calculate the FPKM values. Pearson and Spearman correlation coefficient ($r$) and correlation $P$ value (p) were used to assess the co-expression relationship by an in-house R script.$r > 0.8$ or $r < -0.8$ with pvalue < 0.01 was considered as strong correlation, and $r \geq 0.9$ or $r \leq -0.9$ with p value <0.01 was regarded as strikingly correlation.

### *Statistical analysis*

In-house Perl scripts were used for calculating length and GC contents of LLM PCGs and lncRNAs. One-way ANOVA was used for pairwise comparisons. For all the comparison, * and ** indicate $p < 0.05$ and $p < 0.01$, respectively

## Acknowledgments

We would like to acknowledge all the members in Kong laboratory for their helpful discussion and comments.

## Author contributions

KQP and JJJ conceived and designed the study. KQP and JJJ performed all analysis and drafted the manuscript. KQP and JJJ revised the manuscript. All authors read and approved the final manuscript.

## Disclosure statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Funding

This work was supported jointly and equally by grants from the National Key R&D Program of China [No. 2018YFC2000400, 2018YFE0203700]. Additional support was received from the Key Research Program of Frontiers Science [QYZDB-SSW-SMC020] and Youth Innovation Promotion Association [to Q.P.K.] of the Chinese Academy of Sciences, National Natural Science Foundation of China [91749109, 81671404], and Yunnan Applied Basic Research Project [2017FA038].

## References

[1] Seluanov A, Gladyshev VN, Vijg J, et al. Mechanisms of cancer resistance in long-lived mammals. Nat Rev Cancer. 2018;18:433–441.
[2] Azpurua J, Seluanov A. Long-lived cancer-resistant rodents as new model species for cancer research. Front Genet. 2013;3:319.
[3] Mitchell SJ, Scheibye-Knudsen M, Longo DL, et al. Animal models of aging research: implications for human aging and age-related diseases. Annu Rev Anim Biosci. 2015;3:283–303.
[4] Albuquerque TAF, Val LDD, Doherty A, et al. From humans to hydra: patterns of cancer across the tree of life. Biol Rev. 2018;93:1715–1734.
[5] Szymanska H, Lechowska-Piskorowska J, Krysiak E, et al. Neoplastic and nonneoplastic lesions in aging mice of unique and common inbred strains contribution to modeling of human neoplastic diseases. Vet Pathol. 2014;51(3):663–679.
[6] Lipman R, Galecki A, Burke DT, et al. Genetic loci that influence cause of death in a heterogeneous mouse stock. J Gerontol Ser A. 2004;59:B977–83.
[7] Cancer Statistics. National Cancer Institute. 2015. Available from: https://www.cancer.gov/about-cancer/understanding/statistics
[8] Seluanov A, Hine C, Azpurua J, et al. Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. Proc Natl Acad Sci. 2009;106:19352–19357.
[9] Tian X, Azpurua J, Hine C, et al. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. Nature. 2013;499:346–349.
[10] Wilkinson GS, South JM. Life history, ecology and longevity in bats. Aging Cell. 2002;1:124–131.
[11] Seim I, Fang X, Xiong Z, et al. Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*. Nat Commun. 2013;4:1–8.
[12] Huang Z, Jebb D, Teeling EC. Blood miRNomes and transcriptomes reveal novel longevity mechanisms in the long-lived bat, Myotis myotis. BMC Genomics. 2016;17:906.
[13] Zhang G, Cowled C, Shi Z, et al. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science. 2013;339:456–460.
[14] Khaidakov M, Siegel ER, Shmookler Reis RJ. Direct repeats in mitochondrial DNA and mammalian lifespan. Mech Ageing Dev. 2006;127:808–812.
[15] Shen -Y-Y, Liang L, Zhu Z-H, et al. Adaptive evolution of energy metabolism genes and the origin of flight in bats. Proc Natl Acad Sci. 2010;107:8666–8671.
[16] Seim I, Ma S, Zhou X, et al. The transcriptome of the bowhead whale *Balaena mysticetus* reveals adaptations of the longest-lived mammal. Aging (Albany NY). 2014;6(10):879.
[17] Keane M, Semeiks J, Webb AE, et al. Insights into the evolution of longevity from the bowhead whale genome. Cell Rep. 2015;10:112–122.
[18] Lin A, Li C, Xing Z, et al. The LINK-A lncRNA activates normoxic HIF1α signalling in triple-negative breast cancer. Nat Cell Biol. 2016;18:213–224.
[19] Hu Q, Ye Y, Chan L-C, et al. Oncogenic lncRNA downregulates cancer cell antigen presentation and intrinsic tumor suppression. Nat Immunol. 2019;20:835–851.
[20] Pickard MR, Williams GT. Molecular and cellular mechanisms of action of tumour suppressor GAS5 LncRNA. Genes (Basel). 2015;6:484–499.
[21] Ji J, Dai X, Yeung S-CJ, et al. The role of long non-coding RNA GAS5 in cancers. Cancer Manag Res. 2019;11:2729–2737.
[22] Wang Y, Xin S, Zhang K, et al. Low GAS5 levels as a predictor of poor survival in patients with lower-grade gliomas. J Oncol. 2019;2019:1–15.
[23] Sun M, Jin F, Xia R, et al. Decreased expression of long noncoding RNA GAS5 indicates a poor prognosis and promotes cell proliferation in gastric cancer. BMC Cancer. 2014;14:319.
[24] Dong S, Zhang X, Liu D. Overexpression of long noncoding RNA GAS5 suppresses tumorigenesis and development of gastric cancer by sponging miR-106a-5p through the Akt/mTOR pathway. Biol Open. 2019;8(6):bio041343.
[25] Wen Q, Liu Y, Lyu H, et al. Long noncoding RNA GAS5, which acts as a tumor suppressor via microRNA 21, regulates cisplatin resistance expression in cervical cancer. Int J Gynecol Cancer. 2017;27:1096–1108.
[26] Ye K, Wang S, Zhang H, et al. Long noncoding RNA GAS5 suppresses cell growth and epithelial–mesenchymal transition in osteosarcoma by regulating the miR-221/ARHI pathway. J Cell Biochem. 2017;118:4772–4781.
[27] Jiang JJ, Cheng LH, Wu H, et al. Insights into long noncoding RNAs of naked mole rat (*Heterocephalus glaber*) and their potential association with cancer resistance. Epigenetics Chromatin. 2016;9(1):51.
[28] Yim H-S, Cho YS, Guang X, et al. Minke whale genome and aquatic adaptation in cetaceans. Nat Genet. 2014;46:88–92.
[29] Kim EB, Fang X, Fushan AA, et al. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. Nature. 2011;479:223–227.
[30] Li L, Stoeckert CJ, Roos DS. OrthoMCL: Identification Of Ortholog Groups For Eukaryotic Genomes. Genome Res. 2003;13:2178–2189.
[31] Turbill C, Bieber C, Ruf T. Hibernation is associated with increased survival and the evolution of slow life histories among mammals. Proc R Soc B Biol Sci. 2011;278:3355–3363.
[32] Turbill C, Prior S. Thermal climate-linked variation in annual survival rate of hibernating rodents: shorter winter dormancy and lower survival in warmer climates. Funct Ecol. 2016;30:1366–1372.
[33] Zhao M, Kim P, Mitra R, et al. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. Nucleic Acids Res. 2016;44:D1023–31.

[34] Vazquez JM, Sulak M, Chigurupati S, et al. A zombie LIF gene in elephants is upregulated by TP53 to induce apoptosis in response to DNA damage. Cell Rep. 2018;24:1765–1776.
[35] Kirk JM, Kim SO, Inoue K, et al. Functional classification of long non-coding RNAs by k-mer content. Nat Genet. 2018;50:1474–1482.
[36] Lagunas-Rangel FA. Why do bats live so long? —Possible molecular mechanisms. Biogerontology.2020 Feb;21(1):1–1.
[37] Buffenstein R. The naked mole-rat: a new long-living model for human aging research. J Gerontol A Biol Sci Med Sci. 2005;60:1369–1377.
[38] Kour S, Rath PC. Long noncoding RNAs in aging and age-related diseases. Ageing Res Rev. 2016;26:1–21.
[39] Kim NW, Piatyszek MA, Prowse KR, et al. Specific association of human telomerase activity with immortal cells and cancer. Science. 1994;266:2011–2015.
[40] Buseman CM, Wright WE, Shay JW. Is telomerase a viable target in cancer? Mutat Res. 2012;730:90–97.
[41] Raveh E, Matouk IJ, Gilon M, et al. The H19 Long non-coding RNA in cancer initiation, progression and metastasis – a proposed unifying theory. Mol Cancer. 2015;14:184.
[42] Sulak M, Fong L, Mika K, et al. TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. eLife. 2016;5:e11994.
[43] Blagosklonny MV. Big mice die young but large animals live longer. Aging (Albany NY). 2013;5:227–233.
[44] Lin C, Yang L. Long noncoding RNA in cancer: wiring signaling circuitry. Trends Cell Biol. 2018;28:287–301.
[45] Carlevaro-Fita J, Lanzós A, Feuerbach L, et al. Cancer LncRNA Census reveals evidence for deep functional conservation of long noncoding RNAs in tumorigenesis. Commun Biol. 2020;3(1):1–16.
[46] Diederichs S. The four dimensions of noncoding RNA conservation. Trends Genet. 2014;30(4):121–123.
[47] Marín-Béjar O, Mas AM, González J, et al. The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. Genome Biol. 2017;18(1):202.
[48] Cox MP, Peterson DA, Biggs PJ. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics. 2010;11:485.
[49] Kim D, Paggi JM, Park C, et al. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37:907–915.
[50] Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562–578.
[51] Kang Y-J, Yang D-C, Kong L, et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. 2017;45:W12–6.